Niveau(x) d'enseignement: 1 Numéro EDUQ : 8670199
Centre d'intérêt: 640

Makdissi, Marie, 1940- .

Evaluation en contexte scolaire d'un logiciel pour l'administration automatisée de tests / par Marie Makdissi. -- Sainte-Foy : INRS-Education, 1984.
18, [12] f. ; 28 cm.

"Communication présentée au Congrès 1984 de l'Association internationale de pédagogie expérimentale de langue française (A.I.P.E.L.F.) Sherbrooke, Québec". Code: C-123. Bibliogr.: f. 18.

Vedette(s)-secondaire(s): INRS-Education

Descripteur(s): Elaboration de logiciels * Evaluation de logiciels * Application pédagogique de l'ordinateur * Elaboration de tests * Evaluation de l'apprentissage * Enseignement géré par ordinateur * Primaire (Ordre d'enseignement) * Secondaire (Ordre d'enseignement)

Dans cet exposé, l'auteure décrit un logiciel d'administration automatisé des tests. Elle retrace, dans un premier temps, le contexte dans lequel le logiciel a été conçu, en donne une brève description, puis présente successivement le plan d'évaluation qui fut arrêté et sa réalisation au cours de l'année 1983-1984, ainsi que les résultats de cette évaluation. Elle énumère, par la suite, les améliorations apportées au logiciel dans sa version finale, améliorations qui furent dictées par les données de l'observation ou encore par la réflexion des membres de l'équipe de recherche de l'INRS-Education. Elle traite, en conclusion, succinctement, de l'avenir immédiat de ce logiciel.

Type de document: r

EDUQ 8670199

Marie Makdissi

EVALUATION EN CONTEXTE SCOLAIRE D'UN LOGICIEL POUR L'ADMINISTRATION AUTOMATISEE DE TESTS.

INRS-Education

C-123

1984

EDUQ 8670199

Évaluation en contexte scolaire d'un logiciel pour l'administration automatisée de tests

par

Marie Makdissi

Communication présentée au Congrès 1984
de l'Association internationale de
pédagogie expérimentale de langue française
(A.I.P.E.L.F.)

Sherbrooke
Québec

C-123

INRS-Éducation
2383, Chemin Sainte-Foy
Sainte-Foy, Québec
G1V 1T1

Les nombreux discours qui se tiennent depuis quelques années sur l'utilisation de l'ordinateur personnel en classe peuvent souvent donner l'impression que la micro-informatique est un acquis pour l'école. La réalité est toute autre. Il suffit d'entendre les propos des enseignants[1] pour réaliser la diversité de leurs attitudes face à l'implantation de cette nouvelle technologie. Pour les pessimistes, la micro-informatique est une sorte de cataclysme qui vient ébranler leur croyance profonde car elle est perçue comme une rivale qui s'appropriera les tâches d'enseignement; pour les opportunistes, par contre, elle renferme des vertus magiques qui guériront les maux de tête qu'occasionnent trop souvent la préparation de cours et la correction de travaux ou d'examens; enfin, pour la majorité, la micro-informatique est objet d'interrogation ou d'anxiété car son utilisation nécessitera une organisation particulière de la classe et des ajustements pédagogiques, la création de didacticiels de qualité dont le contenu devra s'intégrer harmonieusement aux programmes d'étude.

Conscients des nombreux problèmes suscités par une implantation plus ou moins massive de l'ordinateur personnel en classe et désireux d'y apporter certains éléments de réflexion, une équipe de l'INRS-Éducation[2] a mis sur pied un projet de recherche[3] visant à évaluer la pertinence de l'utilisation par l'enseignant ou l'enseignante d'un ordinateur personnel pour fin de vérification des apprentissages centrés sur les objectifs. Notre recherche poursuivait deux buts importants: concevoir et évaluer un système pour l'administration automatisée de tests

---

1. Dans cette communication, le mot "enseignants" réfère tant aux enseignantes qu'aux enseignants et celui "d'étudiants", tant aux étudiantes qu'aux étudiants.
2. Cette équipe de l'INRS-Éducation est composée des personnes suivantes: Yves Bégin et Denis Rhéaume, professeurs-chercheurs; Marie Makdissi, agente de recherche; Normand Dufour, assistant de recherche; Marie-Marthe Hébert et Josée Brouard, stagiaires.
3. Ce projet est subventionné par le Fonds pour la formation de chercheurs et l'action concertée du Québec (FCAC-EQ48).

et, parallèlement, concevoir et évaluer un système automatisé pour la gestion des dossiers des étudiants en fonction des objectifs spécifiques d'apprentissage.

Nous nous limiterons, dans cet exposé, au logiciel d'administration automatisée de tests. Après avoir décrit le contexte dans lequel a été conçu le logiciel et en avoir donné une brève description, nous présenterons successivement le plan d'évaluation que nous avons arrêté et sa réalisation au cours de l'année 1983-1984, ainsi que les résultats de cette évaluation. Par la suite, nous énumérerons les améliorations apportées au logiciel dans sa version finale, améliorations dictées par les données de l'observation ou encore par la réflexion des membres de l'équipe de recherche. Enfin, nous parlerons succinctement de l'avenir immédiat du logiciel.

CONTEXTE DANS LEQUEL A ÉTÉ CONÇU LE LOGICIEL

Deux raisons principales expliquent le choix que nous avons fait de concentrer nos efforts sur l'ordinateur, instrument de gestion pédagogique. La première appartient à l'histoire de l'INRS-Éducation ou relève de notre credo pédagogique. En effet, depuis 1973, plusieurs chercheurs du centre se sont concertés pour développer un système d'enseignement individualisé connu sous le sigle de SAGE (Système d'Apprentissage Géré par l'Étudiant). Dans un tel contexte, deux problèmes se posent immédiatement à l'enseignant ou à l'enseignante: celui de la gestion des objectifs ou de la gestion du dossier des élèves, chacun développant son propre programme d'étude, et celui de l'administration d'une multiplicité des tests. Quant à la deuxième raison explicative de notre option concernant l'ordinateur gestionnaire, elle nous a été dictée par la prudence qui s'impose face à l'envahissement d'une technologie nouvelle. Nous n'avons pas voulu entreprendre l'aventure de l'utilisation de l'ordinateur personnel à des fins d'enseignement de contenu. L'enseignement

par ordinateur est encore très difficile et sera toujours très exigeant quant à la qualité pédagogique; un didacticiel valable doit renfermer en somme toutes les vertus d'un bon pédagogue. C'est pourquoi, nous avons pensé que l'ordinateur personnel doit d'abord faire ses preuves, comme outil, dans des tâches plus modestes au départ et s'intégrer ainsi plus facilement à la vie de la classe. D'ailleurs, comme le souligne Decker F. Walker dans un de ses articles sur les limites et les possibilités de l'ordinateur personnel, "pour au moins la durée de la prochaine décennie, il ne faut pas s'attendre à ce que l'ordinateur puisse assumer un rôle important dans les formes conventionnelles de l'éducation".[4]

DESCRIPTION DU LOGICIEL SOUMIS À L'ÉVALUATION

Globalement, le logiciel d'administration de tests, conçu et développé à l'INRS-Éducation au cours de l'année 1982-1983 et évalué en contexte scolaire pendant l'année 1983-1984, se veut un outil qui permette à l'enseignant(e) de construire une banque d'items à partir de laquelle sont générés des tests composés d'items choisis au hasard. Ces tests peuvent être imprimés ou encore, présentés à l'écran directement à l'étudiant qui se sert, à ce moment-là, du terminal pour y répondre.

D'une façon plus précise, le logiciel d'administration automatisée de tests possède quatre caractéristiques importantes pour lesquelles ont opté les membres de l'équipe de recherche chargée de sa conception.

La première particularité du système réside dans la possibilité offerte de générer un test basé sur un échantillon stratifié d'items. Cette stratification repose essentiellement sur un procédé de regroupement des items d'une banque c'est-à-dire sur la formation de sous-ensembles d'items identifiés dans le logiciel par le mot "catégories". Une

---

4. Walker, Decker F., Reflections on the Educational Potential and Limitations of Microcomputers. Phi Delta Kappan, vol. 65, no 2, 1983, pp. 103-107.

catégorie représente un groupe d'items qui réfèrent à un noyau commun; ce noyau peut correspondre soit à une section précise d'une matière enseignée, soit à un objectif spécifique d'apprentissage ou encore, soit à tout autre classement jugé pertinent par l'utilisateur. De plus, l'importance respective accordée à chacune des catégories lors de la génération d'un test peut varier indéfiniment; l'utilisateur, en effet, peut à tout moment changer et déterminer le nombre d'items de chacune des catégories qu'il veut retrouver dans un test.

Une deuxième particularité intéressante du logiciel d'administration automatisée de tests se traduit dans la diversité des types d'items. Un même test peut être composé de cinq types différents d'items, cinq étant le maximum qu'offre le système. Le type d'un item est caractéristique de la réponse que doit fournir l'étudiant à une question posée; il réfère précisément à une façon particulière de répondre à une question.

La variété des types d'items correspond donc à la diversité des types de questions et de réponses que le système peut emmagasiner et traiter. Référant à la diversité des types de réponses que l'étudiant peut être amené à fournir, on peut énumérer ainsi les cinq types d'items ou de questions possibles: la question purement objective où l'étudiant doit indiquer par un "vrai" ou par un "faux" son opinion sur l'exactitude de l'énoncé, la question à choix multiple de réponses, la question dont la réponse ne peut être qu'un seul mot ou un seul nombre, la question dont la réponse peut comporter une chaîne de mots et enfin la question exigeant comme réponse un estimé numérique.

La troisième caractéristique importante du logiciel réfère à la possibilité offerte à l'utilisateur d'inclure dans une banque des items dont la coapparition dans un même test est jugée inacceptable. Lors de la génération d'un test, le choix d'un de ces items empêche automatiquement la présence des autres. Ainsi les problèmes de validité engendrés

par la coapparition de certains items dans un même test se trouvent solutionnés. En effet, il peut arriver que la présence d'un item dans un test contienne "un indice important de la réponse à un autre item" ou encore "que l'on désire placer dans une même banque deux items de forme différente mais dont la similitude de contenu rend injustifiable leur coapparition".[5] Cette particularité du logiciel d'empêcher dans un test la coapparition d'items jugés incompatibles permet de conserver dans une même banque plusieurs items dont la valeur ne peut être contestée que s'ils sont générés simultanément.

La quatrième caractéristique du logiciel, propre à la situation d'évaluation interactive, réside dans la souplesse offerte à l'étudiant lors de la passation d'un test. L'étudiant, en effet, répond à tous les items d'un test avant de voir ses réponses corrigées; il peut même revenir sur des items particuliers ou encore réviser son test avant d'en demander la correction. Le "feed-back" à chacune des questions n'est fourni qu'au terme de l'activité d'évaluation, au moment où l'étudiant décide d'y mettre fin.

Les raisons explicatives des options que nous avons faites concernant les caractéristiques du logiciel d'administration automatisée de tests se retrouvent en partie consignées dans un document de l'INRS-Éducation intitulé: "L'utilisation dans un contexte d'évaluation formative d'un test généré par ordinateur à partir d'une banque d'items".[6]

PLAN D'ÉVALUATION

Si la prudence nous a guidés vers la conception de logiciels centrés sur la vérification des apprentissages plutôt que vers la création

---

5. Rhéaume, D., Dufour, N., Makdissi, M., Bégin, Y. L'utilisation dans un contexte d'évaluation formative d'un test généré par ordinateur à partir d'une banque d'items. INRS-Éducation, C-110, 1983.
6. Idem.

de didacticiels proprement dits, elle nous a aussi amenés à considérer comme essentielle une vérification systématique de nos logiciels dans le but de bien en cerner les avantages et les limites avant de les offrir aux enseignants.

Le système a été soumis à une pré-évaluation en laboratoire au printemps et à l'été 1983. En plus de nous donner l'occasion d'identifier et de corriger certaines lacunes inhérentes au système, cette pré-évaluation nous a confirmé que le logiciel était assez fiable pour être mis à la disposition d'enseignants.

En ce qui concerne l'évaluation proprement dite, nous avons retenu cinq critères pour l'effectuer. Ces critères sont les suivants:

1. le nombre, l'importance et l'effet des défectuosités techniques rencontrées durant la période d'évaluation;
2. le degré d'utilisation de chacune des fonctions du logiciel;
3. les avantages perçus par l'enseignant ou l'enseignante dans l'utilisation du logiciel;
4. les réactions des étudiants ou des étudiantes;
5. le degré d'appropriation par l'enseignant ou l'enseignante du logiciel de gestion.

RÉALISATION DU PLAN D'ÉVALUATION

L'évaluation du logiciel d'administration automatisée de test a été amorcée en septembre 1983. La première étape de l'année scolaire a permis de recruter deux enseignants, de les sensibiliser à l'utilisation du système et de leur apporter le soutien nécessaire à la construction d'une banque d'items. La période d'évaluation proprement dite a débuté en janvier 1984 pour se poursuivre jusqu'au début d'avril 1984. Les étudiants de deux écoles différentes, de niveau secondaire, ont utilisé

le logiciel. Il s'agit des 46 étudiantes de secondaire III, option latin, de l'École secondaire des Ursulines de Québec et de 20 étudiants de secondaire V, option commerce, de l'École secondaire Marie-de-l'Incarnation[7].

Les classes choisies ont reçu la visite, au moins hebdomadaire, d'une observatrice de l'INRS-Éducation chargée de recueillir les données pertinentes à l'évaluation et de procurer, au besoin, le support nécessaire aux utilisateurs. De plus, chacune des deux observatrices tenaient un journal faisant état de sa participation à la vie de classe ou encore de sa perception des attitudes et des réactions de l'enseignant et des étudiants face à l'utilisation de l'ordinateur personnel. Les deux observatrices sont des étudiantes universitaires, l'une de troisième cycle et l'autre de deuxième cycle, engagées comme stagiaires à l'INRS-Éducation, dans le but de s'initier à la recherche.

Comment a-t-on procédé à l'évaluation de chacun des cinq critères cités précédemment?

L'évaluation du premier critère a consisté pour l'observatrice à noter, lors de ses visites en classe, les problèmes techniques rencontrés et à s'informer auprès de l'enseignant des problèmes qui ont surgi en son absence. De plus, chacun des étudiants complétait, après une période d'utilisation du système, une grille d'observations conçue spécialement pour eux (cf. Annexes I et II).

Quant aux critères concernant l'utilisation des fonctions du système, les avantages et les limites de l'instrument et le degré d'appropriation par l'enseignant du système, ils ont été, bien sûr, présents à l'esprit des observatrices qui notaient les commentaires ou les réactions dont elles étaient témoins. Cependant, l'information touchant ces

---

7. Les deux enseignants responsables de ces groupes sont Monsieur André Doyon et Madame Lucille Paquet.

trois critères a été recueillie principalement par l'intermédiaire d'un questionnaire qui a été élaboré par les membres de l'INRS-Éducation (cf. Annexe III) et complété par les enseignants à la fin de la période d'observation dans les classes, soit au début du mois d'avril 1984.

Enfin, le critère relatif aux réactions des étudiants, en plus d'avoir reçu une attention particulière de la part des observatrices, a été évalué dans un cas par l'intermédiaire de la grille d'observation qui était destinée aux étudiantes et qui accordait une place à l'appréciation personnelle de chacune (cf. Annexe I) et, dans l'autre cas, par le recours, à la fin de la période d'observation, à une plénière où les étudiants ont livré leur perception de l'utilisation du système.

RÉSULTATS DE L'ÉVALUATION

Les trois mois d'expérimentation et d'observation-participante nous ont conduits dans l'ensemble à une évaluation positive du logiciel d'administration automatisée de tests, compte tenu des cinq critères d'évaluation retenus.

Aucune défectuosité technique majeure n'est venue entraver l'expérimentation; cela est dû, croyons-nous, en grande partie à la période intense d'utilisation des deux premiers mois faite à l'INRS-Éducation lors de l'initiation des enseignants. Pendant cette période, en effet, plus de 300 items ont été entrés dans le système et de nombreuses simulations de passation de tests au terminal ont été menées. Cette pré-expérimentation en laboratoire nous a permis, en un laps de temps très court, d'exploiter au maximum les différentes fonctions du système et d'identifier très rapidement les faiblesses techniques. Cependant, le guide d'utilisation destiné à l'enseignant n'ayant été rédigé formellement qu'en juin et juillet 1984, certains problèmes techniques mineurs ont été rencontrés. Ces difficultés par contre relevaient surtout de la

nouveauté de l'instrument utilisé et du manque de familiarisation de la part des enseignants et des étudiants avec les appareils informatiques. La connaissance des diverses fonctions, des possibilités et des limites du système une fois maîtrisée par les enseignants et les étudiants, l'expérimentation nous a montré que le logiciel ne recelait pas de difficultés techniques majeures.

Les principales fonctions du système ont été utilisées proportionnellement à ce que nous avions prévu. Aucune fonction n'a été négligée ou considérée inutile par les enseignants, même si certaines ont été privilégiées. Toutefois, l'expérimentation nous a permis de rendre plus limpides certaines fonctions du logiciel, celles en particulier reliées à la passation de tests; elle nous a amenés aussi à augmenter les possibilités de certaines opérations et à n'en ajouter de nouvelles. Nous en traiterons dans la partie suivante de notre exposé intitulée "les améliorations apportées au logiciel".

D'aucuns pourraient nous reprocher la limite de notre évaluation qui résulte du fait que l'équipe de conception du logiciel a été sensiblement la même que celle chargée de son évaluation. Nous en étions conscients et le questionnaire adressé aux enseignants a été pour nous un moyen pour contourner partiellement l'écueil de notre subjectivité; il a été aussi un moyen pour évaluer les deux critères concernant respectivement les avantages et les désavantages de l'utilisation du logiciel tels que perçus par les enseignants, le degré d'appropriation du logiciel ressenti par les enseignants.

Les deux enseignants qui ont expérimenté le logiciel d'administration automatisée de tests ont été unanimes à reconnaître la valeur de cet instrument dans le cadre de l'évaluation formative. Pour eux, en effet, cet outil a rendu possible le développement d'habitudes d'autoévaluation chez les étudiants; il a de plus initier progressivement les étudiants à une lecture attentive des questions posées et à la rigueur des réponses données.

Le plus grand avantage reconnu par les utilisateurs a été la souplesse inhérente au logiciel pour la production de tests variés respectant l'importance accordée à l'une ou l'autre des sections de la matière enseignée. En effet, les deux enseignants ont eu constamment recours, lors de la construction des tests, à la possibilité de l'échantillon stratifié d'items, la possibilité de la génération des items au hasard pour l'ensemble de la banque n'étant à peu près pas utilisée sinon pour certains tests dits de révision. Les raisons évoquées pour expliquer leur nette préférence pour l'échantillon stratifié d'items sont les suivantes. Premièrement, l'importance d'une catégorie (ou groupe) d'items n'est pas directement reliée au nombre d'items disponibles dans la banque pour cette catégorie. En second lieu, l'importance accordée à une catégorie d'items peut varier presque à l'infini sans l'obligation d'ajouter ou de retrancher des items dans la banque. Troisièmement, la prise en compte par le système d'items jugés équivalents et enfin, la qualité de la représentativité de l'ensemble des items d'un test font du logiciel, à leur avis, un outil efficace pour élaborer des tests d'évaluation, assurer la passation des tests ainsi que la correction.

Nous ne voudrions pas laisser sous-entendre que l'évaluation du logiciel faite par les deux enseignants s'est limitée à l'acceptation totale du produit que nous avons mis à leur disposition. Au contraire, leurs commentaires et leurs suggestions ont été mis à profit pour nous permettre d'apporter les améliorations nécessaires à un meilleur rendement du système. Cependant, certains souhaits émis, comme par exemple l'ajout de types d'items, n'ont pu être retenus à cause des capacités mêmes de l'équipement utilisé (ordinateur personnel Apple II de 48 K, un lecteur de disquette conventionnel, un écran de 40 colonnes). De plus, certaines limites, comme la capacité maximale d'une banque limitée à 160 items ou encore la taille maximale d'un test ne pouvant excéder 40 items, n'ont pu être enlevées compte tenu du choix que nous avions fait concernant les caractéristiques essentielles du logiciel.

En ce qui concerne le critère relatif au degré d'appropriation par l'enseignant du logiciel d'administration automatisée de tests, nous pouvons conclure qu'il est directement relié à la compréhension du système et particulièrement à l'opération de construction de la banque. Si nous nous référons tant aux avis des deux utilisateurs qu'aux rapports des deux observatrices de l'INRS-Éducation, la création d'items est l'opération la plus exigeante car elle nécessite une réflexion soutenue concernant la formulation particulière des énoncés et des paramètres. Des questions comme, par exemples, celles concernant l'exactitude des réponses attendues, la latitude permise lors de la correction ou encore le choix du type d'items en fonction de l'énoncé sont autant d'interrogations que l'utilisateur doit avoir en tête au moment où il construit une banque. De plus, l'impossibilité dans certains cas de prévoir toutes les réponses valables lors de l'élaboration des clés de correction oblige l'enseignant à "se soumettre à de fréquentes modifications essentielles à l'amélioration de sa banque"[8].

Le haut degré d'appropriation du logiciel manifesté par les deux enseignants pendant l'expérimentation a été proportionnel à la qualité de leur travail professionnel. Si le système possède un degré de transparence qui en facilite l'appropriation, il nous faut admettre, de concert avec les deux enseignants impliqués dans l'expérience et les deux observatrices, qu'il est nécessaire que l'utilisateur accorde "un bon temps de préparation axé sur le contenu des questions, des réponses et des commentaires ajoutés aux réponses... C'est leur adéquation à la matière et à la clientèle, leur variété, leur précision, même leur nouveauté... qui feront de ce logiciel un instrument"[9] favorisant l'évaluation formative des apprentissages.

---

8. Brouard, J., Observation-participante de l'utilisation dans une classe du logiciel "Construction d'une banque d'items et passation de tests au terminal" à l'École secondaire Marie-de-l'Incarnation. INRS-Éducation, R-172, 1984.
9. Hébert, M.M., Observation-participante de l'utilisation dans une classe du logiciel "Construction d'une banque d'items et passation de tests au terminal" à l'École secondaire des Ursulines de Québec. INRS-Éducation, R-173, 1984.

Quant aux réactions des étudiants face à l'utilisation du logiciel, nous disposons pour les évaluer des commentaires des observatrices et des enseignants ainsi que des avis des étudiants formulés soit lors de la plénière destinée à recueillir leurs opinions, soit sur la partie "appréciation" de la grille d'observations (cf. Annexe I). Ce critère est évidemment retenu en fonction de la situation interactive d'évaluation, c'est-à-dire en regard des réactions d'étudiants pendant ou après une session d'évaluation.

Dans l'ensemble, les réactions des étudiants ont été très favorables même après l'excitation momentanée qu'occasionne actuellement l'installation d'un ordinateur personnel dans une école qui n'en possède pas encore. Les raisons évoquées par les étudiants et explicatives de leur satisfaction face à l'utilisation du logiciel sont de taille. Pour une très grande majorité d'entre eux, la "passation régulière de tests au terminal leur permet de mieux prendre conscience de ce qu'ils savent ou ne savent pas d'un chapitre ou d'une notion en particulier"[10]. Ce genre d'évaluation est l'occasion d'approfondir une matière de "façon moins ennuyeuse que l'écoute ou la lecture"[11]; ce contexte interactif d'évaluation a le mérite, selon les étudiants des deux classes expérimentales, de favoriser l'apprentissage de notions ou leur consolidation. Et bien que "la sévérité de correction soit parfois considérée comme un désagrément, les élèves reconnaissent que cela peut les aider grandement dans leur maîtrise de la matière et de l'orthographe en particulier"[12].

Les résultats de l'évaluation en fonction des cinq critères que nous avions retenus confirment, à notre avis, la pertinence de l'utilisation par l'enseignant et ses élèves d'un système conçu pour l'administration automatisée de tests. Compte tenu de ces résultats, l'équipe de

---

10. Brouard, J., op. cit.
11. Idem.
12. Idem.

recherche chargée de la conception et de l'évaluation du logiciel a procédé au terme de la période expérimentale à la finalisation de l'instrument.

VERSION FINALE DU LOGICIEL

Nous soulignons immédiatement que le logiciel d'administration automatisée de tests est connu, depuis juin 1984 sous le nom de "Rétroaction". De plus, un guide d'utilisation[13] est disponible depuis la fin de juillet 1984, guide dont le but est de faciliter la familiarisation des futurs utilisateurs et où le lecteur peut retrouver rapidement, dès les toutes premières pages, une description claire et rapide du logiciel. Nous nous limiterons ici à souligner les principales améliorations qui ont été apportées au logiciel à la lumière de l'expérimentation.

À la demande expresse des enseignants (cela correspondait d'ailleurs à un voeu de l'équipe de recherche), une ouverture a été apportée à la génération des tests. Dans sa phase expérimentale, le logiciel offrait deux façons de générer un test: le mode général qui correspondait à un échantillon d'items choisis au hasard à l'intérieur d'une banque ou encore le mode stratifié basé sur l'échantillon stratifié d'items, la stratification ayant trait aux catégories d'items ou groupes d'items. La version finale du logiciel offre un troisième mode de génération de tests, un mode fixe où l'enseignant peut choisir lui-même les items devant faire partie d'un test.

La place accordée au feed-back a été augmentée. Alors que la version expérimentale n'y accordait que 30 caractères, la version finale

---

13. Rhéaume, D., Brouard, J., Guide d'utilisation du logiciel éducatif "Rétroaction". INRS-Éducation, juillet 1984.

offre une possibilité de feed-back pouvant atteindre jusqu'à 5 lignes de 40 caractères dans le cas d'items qui occupent moins d'espace que le maximum prévu. Cette modification est capitale si l'on pense que l'espace réservé au feed-back représente la latitude attribuée à l'enseignant pour fournir, en plus de la bonne réponse, des éléments de contenu ou encore des rappels à des notions antérieures, des références, des explications supplémentaires, etc. Nous notons que la nature du feed-back ne dépend pas de la nature de la réponse donnée par l'élève. C'est plutôt la rétroaction présentée à l'étudiant lorsqu'il passe un test au terminal: le feed-back pour un item est unique et c'est à l'enseignant de l'exploiter de la meilleure façon possible au moment où il construit sa banque d'items. Pour chaque item, le feed-back devrait être le lieu privilégié pour provoquer l'apprentissage.

Une autre amélioration importante apportée au logiciel concerne les deux types de questions exigeant une réponse plus ou moins brève. Dans la version expérimentale, l'utilisateur pouvait uniquement avoir accès à trois clés de correction différentes; dans la version finale, pour augmenter le degré de souplesse de la correction ou encore pour doter le système d'une plus grande flexibilité au niveau de l'énoncé même de l'item, les clés de correction ont été portées à huit. La multiplicité des clés permettra à l'enseignant de prévoir, entre autre, comme bonnes réponses, des termes ou expressions considérés synonymes. La restriction du nombre de clés disponibles avait occasionné certains problèmes lors de la passation des tests; des étudiants se sont vus refuser une bonne réponse parce que le mot qu'ils avaient utilisé, même s'il était juste, n'avait pu être introduit dans les clés disponibles.

Plusieurs modifications mineures ont aussi contribué à l'amélioration de la version finale du logiciel. Pour n'en citer que quelques-unes, soulignons la présence d'un traitement de texte plus raffiné lors de l'entrée des items ou de la présentation générale du test, le changement de certains termes par des termes plus précis et plus clairs, une

lisibilité accrue des écrans en particulier lors de la révision du test par l'étudiant, la fusion de certaines fonctions par souci de simplicité, etc. De plus, des ajouts, comme l'intégration d'un compteur de caractères disponibles lors de l'entrée d'un item, ont été incorporés au logiciel à la demande des enseignants.

La version finale du logiciel d'administration automatisée de tests, identifié depuis la fin de la période expérimentale sous le titre "Rétroaction" représente "une des façons de mettre à contribution l'ordinateur pour l'évaluation des apprentissages scolaires"[14]. La plupart des améliorations souhaitées par les deux enseignants qui l'ont expérimenté font de "Rétroaction" un outil susceptible d'alléger une des tâches de gestion pédagogique les plus significatives dans le cadre du processus apprentissage/enseignement.

AVENIR IMMÉDIAT DU LOGICIEL

Avant de terminer cet exposé, nous tenons à noter que l'expérimentation du logiciel en milieu scolaire a créé un intérêt manifeste chez les enseignants et la direction des deux écoles impliquées dans le processus d'évaluation. À l'école des Ursulines de Québec, la direction de l'institution a réservé deux journées pédagogiques complètes pour sensibiliser ses enseignants aux applications de l'ordinateur personnel; le logiciel d'administration automatisée de tests a servi, à cet égard, d'exemple pour illustrer "une façon heureuse d'introduire l'ordinateur en milieu scolaire"[15]. Pour leur part, plusieurs enseignants ainsi que la direction de l'École Marie-de-l'Incarnation ont manifesté le désir d'obtenir une version "Commodore" du logiciel, convaincus qu'ils étaient des avantages du système et de sa relative simplicité.

---

14. Idem.
15. Idem.

Nous ne pouvons pas présumer de l'avenir du logiciel. Cependant, il importe de souligner l'entente conclue entre l'INRS-Éducation et la Télé-Université concernant la diffusion de "Rétroaction". Dès septembre 1984, la Télé-Université l'incorporera au matériel requis pour le cours d'applications pédagogiques de l'ordinateur personnel auquel devraient s'inscrire plus d'un millier d'enseignants. C'est donc dire que dès le début de la prochaine année scolaire, le logiciel "Rétroaction" connaîtra un rayonnement provincial. Ainsi, les concepteurs du logiciel auront atteint un de leurs principaux objectifs, celui de susciter chez les enseignants le goût de recourir davantage à l'évaluation formative des apprentissages.

CONCLUSION

Après ces deux années consacrées en partie à la réalisation de notre projet sur les applications pédagogiques de l'ordinateur, nous sommes de plus en plus convaincus que l'ordinateur personnel, comme outil de gestion peut remplir de multiples fonctions et particulièrement dans les domaines directement reliés au processus apprentissage/évaluation.

En effet, l'utilisation du logiciel d'administration automatisée de tests en milieu scolaire nous a fourni l'occasion de constater qu'il est possible et relativement facile à des enseignants de favoriser chez les étudiants le développement d'habitudes d'auto-évaluation; le système dans sa forme actuelle s'est avéré être un outil approprié dans un contexte d'évaluation formative... En effet, "l'évaluation formative visant une vérification périodique de la qualité de l'apprentissage au cours du processus d'apprentissage, le système (offre) à l'étudiant une excellente occasion de recevoir rapidement un feed-back et ce, au moment voulu"[16].

---

16. Rhéaume, D., Dufour, N., Makdissi, M., Bégin, Y., L'utilisation dans un contexte d'évaluation formative d'un test généré par ordinateur à partir d'une banque d'items. INRS-Éducation, C-110, 1983.

Les avantages de l'utilisation de l'ordinateur personnel comme outil de gestion nous apparaissent nombreux et indéniables. Si elle permet à l'enseignant d'avoir une meilleure connaissance des acquis de chacun de ses élèves, elle a conséquemment des retombées très positives sur ceux directement impliqués dans le processus d'apprentissage.

BIBLIOGRAPHIE

Brouard, J., Observation-participante de l'utilisation dans une classe du logiciel "Construction d'une banque d'items et passation de tests au terminal" à l'École secondaire Marie-de-l'Incarnation. INRS-Éducation, R-172, 1984.

Hébert, M.M., Observation-participante de l'utilisation dans une classe du logiciel" Construction d'une banque d'items et passation de tests au terminal" à l'École secondaire des Ursulines de Québec. INRS-Éducation, R-173, 1984.

Rhéaume, D., Brouard, J., Guide d'utilisation du logiciel éducatif "Rétroaction". INRS-Éducation, juillet 1984.

Rhéaume, D., Dufour, N., Makdissi, M., Bégin, Y., L'utilisation dans un contexte d'évaluation formative d'un test généré par ordinateur à partir d'une banque d'items. INRS-Éducation, C-110, 1983.

Walker, Decker F., Reflections on the Educational Potential and Limitations of Microcomputers. Phi Delta Kappan, vol. 65, no. 2, 1983, pp. 103-107.

ANNEXE I

OBSERVATIONS PAR L'ELEVE

Nom : _____ Date : _____

Matière : - grammaire latine _____ Heure: _____

- civilisation latine _____ Durée: _____

Résultat du test : _____

DIFFICULTES TECHNIQUES

Si tu as eu des difficultés, indique-les par un crochet, aux lignes ci-dessous :

	crochet ✓
- mise en marche de l'appareil	
- compréhension des directives	
- inscription des réponses	
- possibilité de revoir le test	
- accès aux instructions (CTRL)	
- modification aux réponses inscrites	
- production du corrigé à l'écran	
- correction de l'ordinateur : réponses : bonnes, refusées ..	
- sortie du corrigé sur papier	

- signale tout autre blocage : _____

_____

APPRECIATION

	oui	non
La passation de ce test représente-t-elle pour toi		
- une expérience agréable ?		
- une expérience utile ?		

Si tu as des commentaires, écris-les en quelques mots : _____

_____

_____

ANNEXE II

OBSERVATIONS DE L'ÉLÈVE

Nom : _____ Date : _____

Résultat : _____ % Heure : Début : _____ fin : _____

DIFFICULTÉS RENCONTRÉES

AVANT DE PASSER LE TEST ...

☐ Mise en marche de l'ordinateur : _____
☐ Inscription de ton nom : _____
☐ Autre (s) : _____
_____

AU COURS DE LA PASSATION DU TEST :

☐ Inscription de tes réponses : _____
☐ Modification de tes réponses : _____
☐ Accès aux instructions (CTRL) : _____
☐ Compréhension des directives : _____
☐ Compréhension des questions posées : _____
☐ Révision du test : _____
☐ Modification de tes réponses au cours de la révision : _____
_____
☐ Autre (s) : _____
_____

APRÈS LA PASSATION DU TEST :

☐ Sortie du corrigé sur papier (mise en marche de l'imprimante) : _____
_____
☐ Autre (s) : _____
_____

Si tu as besoin de plus d'espace, tu peux utiliser le verso de cette feuille.

Université du Québec
Institut national de la recherche scientifique

2280, chemin Sainte-Foy
Sainte-Foy, Québec
G1V 1T1
Téléphone : (418) 657-2503

INRS-Éducation

ANNEXE III

Projet A.P.O.
Avril 1984

QUESTIONNAIRE D'ÉVALUATION À L'INTENTION DES ENSEIGNANTS AYANT UTILISÉ LE LOGICIEL DE CONSTRUCTION ET DE PASSATION DE TESTS

A) APPROPRIATION DU LOGICIEL

À l'heure actuelle, avez-vous une bonne connaissance et une bonne maîtrise du logiciel ?

Combien de temps vous a-t-il fallu pour parvenir à ce niveau de connaissance et de maîtrise du logiciel ?

(Indiquez notamment le nombre d'heures)

N. B. N'HÉSITEZ PAS À AJOUTER DES PAGES SUPPLÉMENTAIRES POUR COMPLÉTER VOS RÉPONSES.

Compte tenu des possibilités du logiciel, est-il relativement facile ou relativement difficile d'apprendre à bien l'utiliser ?

Quels sont les aspects qui vous ont semblé les plus difficiles à maîtriser ? Indiquez clairement quels sont ces aspects. Votre réponse pourra éventuellement nous aider à améliorer le logiciel ou le guide de l'utilisateur.

B) AVANTAGES ET INCONVENIENTS DU LOGICIEL

A la lumière de votre expérience d'enseignant(e), quels sont les principaux avantages que vous voyez dans l'utilisation d'un tel logiciel ?

Quels sont les principaux inconvénients que vous voyez dans l'utilisation d'un tel logiciel ?

C) AMÉLIORATIONS SOUHAITABLES DU LOGICIEL

Il s'agit ici d'exprimer vos suggestions quant à l'amélioration à court terme du logiciel. Il se pourrait bien que l'on tienne compte de ces suggestions au moment de mettre au point la version finale du logiciel. Il se pourrait également que certaines suggestions influencent davantage nos travaux futurs.

Suggestions se rapportant à l'entrée des items :

Suggestions se rapportant aux types d'items :

Suggestions se rapportant à l'analyse des réponses par l'ordinateur :

Suggestions se rapportant aux divers tableaux de choix (menus) :

Suggestions se rapportant à la structure d'un test :

Suggestions se rapportant aux catégories et aux restrictions :

Suggestions se rapportant à l'impression d'un test :

Suggestions se rapportant à la présentation d'un test au terminal :

Suggestions se rapportant au rapport imprimé après la passation d'un test :

Autres suggestions d'améliorations :

D) REACTIONS DES ELEVES

La passation d'un test au terminal a-t-elle lieu sans problème ?
Si non, indiquez les principaux problèmes .

Quelles ont été les principales réactions des élèves lors de la passation de tests par l'ordinateur ? Indiquez d'abord et avant tout les réactions qui sont en rapport avec le logiciel .

C) RECOMMANDATION GÉNÉRALE

Considérant la connaissance que vous avez du logiciel et l'utilisation que vous en avez faite durant la période d'évaluation, recommandez-vous à l'équipe de l'INRS-Éducation d'offrir le logiciel à d'autres enseignants intéressés ? (Encerclez le chiffre qui correspond à votre réponse.)

1 Non

2 Oui

3 Oui, à la condition que les changements suivants soient apportés :

(Indiquez clairement ici tous les changements qui doivent être apportés pour que vous recommandiez positivement le logiciel.)

Dans le cas d'une réponse affirmative à la question précédente, considérez-vous que ce logiciel pourrait vous être utile à l'avenir ?

Indiquez les raisons qui justifient votre choix .

G) AUTRES COMMENTAIRES

Auriez-vous d'autres commentaires à faire pour compléter l'évaluation de ce logiciel ?